Arabic Speech Recognition

Marwa F. Mohamed

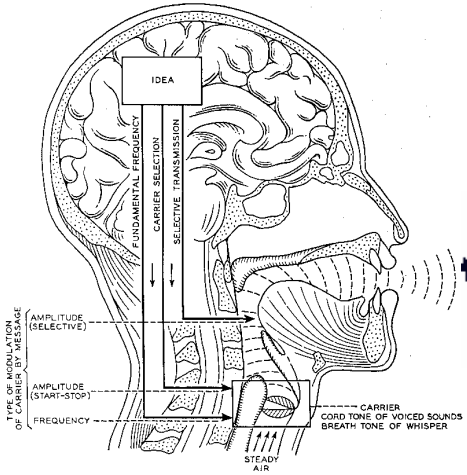
Agenda

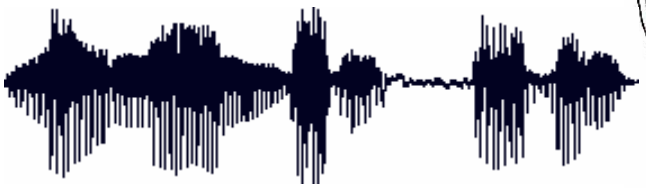
- Automatic Speech Recognition (ASR)
- ASR Engines
- ASR Stages
- ASR Applications
- ASR Challenges

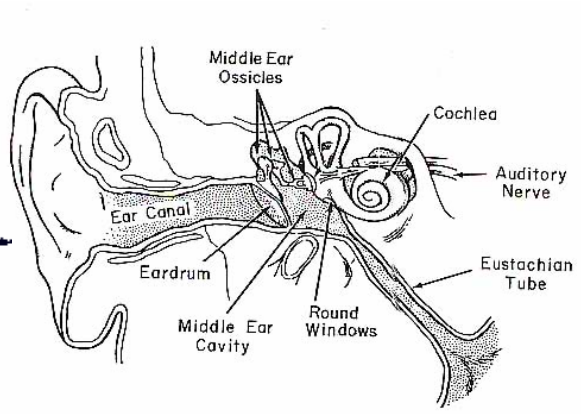
Automatic speech recognition

- Is a branch of artificial intelligence (AI).
- Getting a computer to understand spoken language
- By "understand" we might mean
  - React appropriately
  - Convert the input speech into another medium, e.g. text

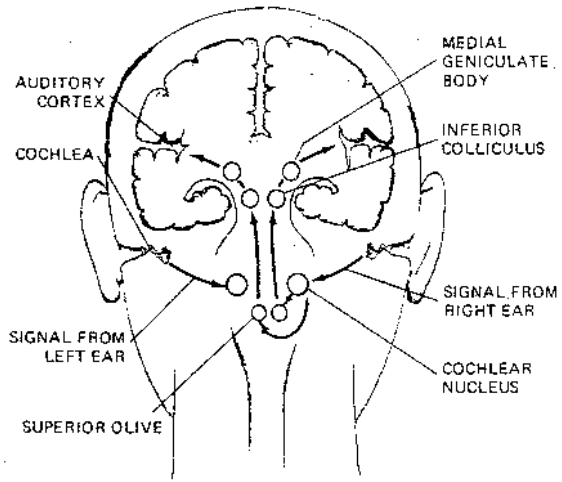
How do humans do it?



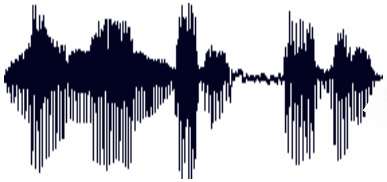




- Articulation produces
- sound waves which
- the ear conveys to the brain
- for processing

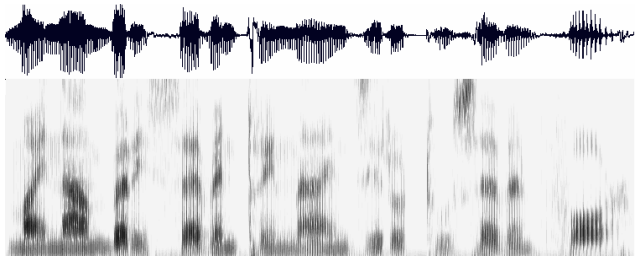


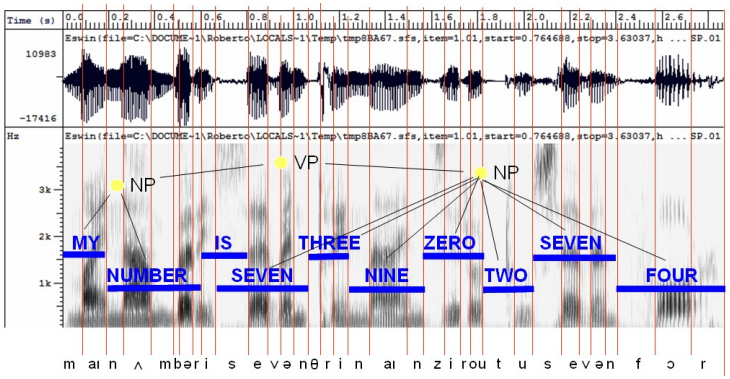
How might computers do it?



Acoustic waveform

Acoustic signal





Speech recognition

- Digitization
- Acoustic analysis of the speech signal
- Linguistic interpretation

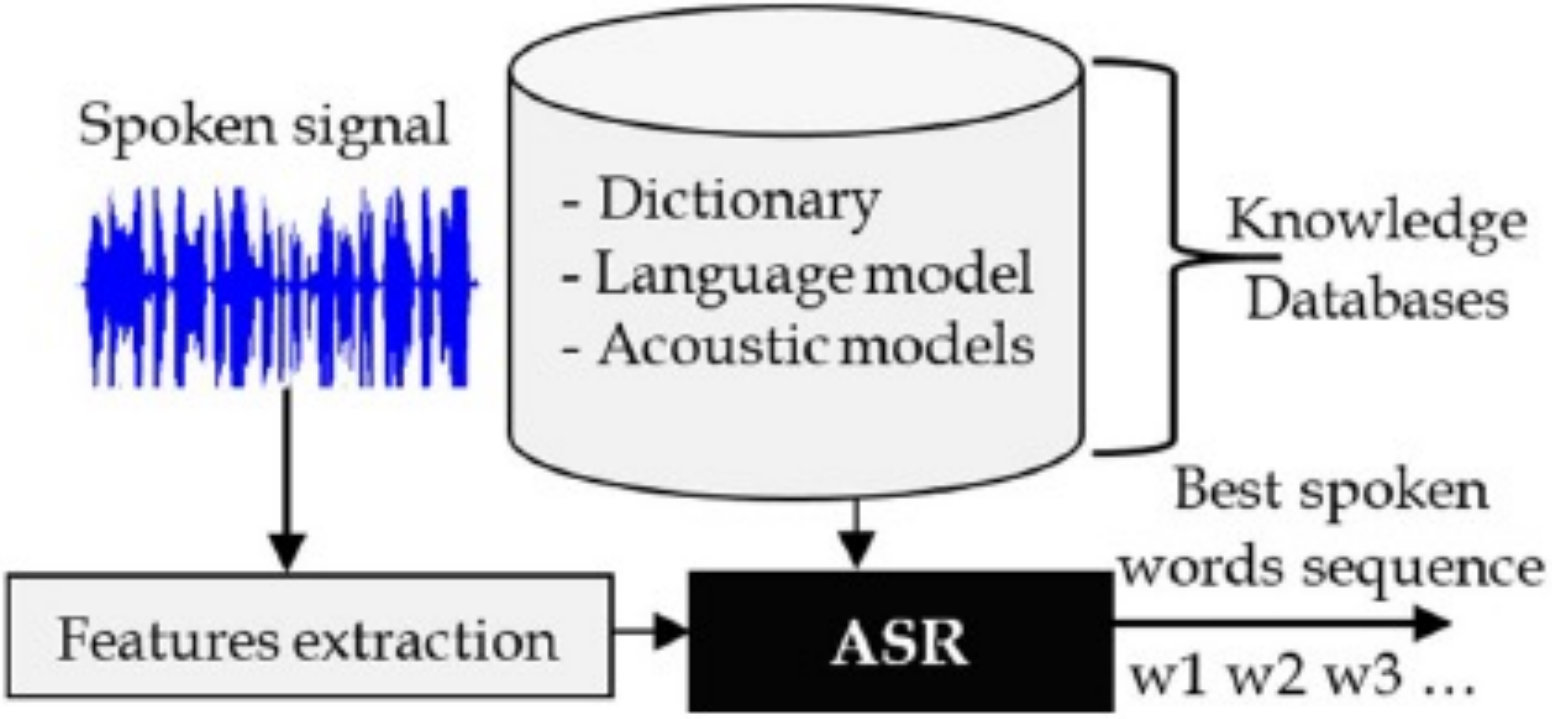
Digitization

- Is the process of converting information into a digital (i.e. computer-readable) format.
- The result is the representation of an object, image, sound, document or signal (usually an analog signal) obtained by generating a series of numbers that describe a discrete set of points or samples.

A standard ASR system consists of

- Microphone unit,
- Speech recognition engine,
- Computer, and
- Certain form of audio/ visual/action output.

ASR Architecture



Arabic Speech Corpora

- The preliminary work in **speech recognition** requires in the first place to specify the type of speech;
  - Isolated-words or
  - Continuous speech.

Arabic Speech Corpora

- In **isolated-words speech** (also called discrete words), a **pause are existed** between digits or words while such constraint is not existed in **continuous or conversational speech**.
- **Isolated-words speech** recognition is characterized by easy to implement when compared to the **continuous speech recognition**.

ASR system stages

- An ASR system comprises **two major stages**:
  - Training stage.
  - Testing stage.

ASR system stages

- **In training stage,** teaching the system by constructing its dictionary and acoustic models for each linguistic unit that the system has to recognize.
- **In testing stage**, the spoken words and all the models available in the dictionary are compared and, in this stage, the system tries to find between them a close match with the slightest mismatch.

ASR system stages

- Feature Extraction and Classification are two principal modules of an **ASR system**.
- Both **testing and training stages involve** feature extraction and classification techniques.
- **Feature Extraction** is the process through which the speech signal is converted into a set of parameters called **speech features**.

ASR system stages

- **Classification task** consists of finding the parameter set from memory which matches as closely as possible the parameter set obtained from the inputted speech signal.
- Commonly **classification techniques** used in ASR framework:
  - Hidden Markov Model (HMM),
  - Artificial neural network (ANN), and
  - Dynamic time warping (DTW)

ASR Applications

- The Applications of an ASR system can be categorized into **two major areas**.
  - In the dictation area.
  - In the human–computer interaction area.

ASR Applications

- In the **dictation area**, the broadcast news dictation technology has been integrated into information extraction and retrieval technologies, many others applications such as
  - Retrieval systems
  - Automatic voice document indexing.
- In the **human–computer interaction area**, a variety of experimental systems for information retrieval through spoken dialogue were explored.

ASR Applications

- Automated conversion of speech into written text, which has the capability to increase output effectiveness and improve access to various computer **applications** such as
  - word processing,
  - remote control using phones,
  - email,
  - speaker identification,
  - language identification

ASR Challenges

- **Arabic** is a Semitic language with approximately 221 million speakers in the Arab world and some African and Asian countries such as Chad, Cyprus, Iran, Kenya, Mali, Niger, Tajikistan, Tanzania, etc. . . .
- There are over 30 different varieties of **colloquial Arabic**.

ASR Challenges

- The absence of unified large continuous speech corpora is an obstacle that might restrain the research in this flourishing domain.

- It has been noticed that almost all-Arabic speech recognition studies have been investigated using **inhouse small corpora.**

Python Code

Speech recognition

- Speech recognition is **a machine's ability to listen to spoken words and identify them**.
- You can then use **speechrecognition** in Python to convert the spoken words into text, make a query or give a reply.

gTTS

- gTTS (Google Text-to-Speech), a Python library and CLI tool to interface with Google Translate's text-to-speech API.
- gTTS is a very easy to use tool which converts the text entered, into audio which can be saved as a mp3 file.
- The gTTS API supports several languages including English, Hindi, Tamil, French, German and many more.
- The speech can be delivered in any one of the two available audio speeds, fast or slow.

```python
from gtts import gTTS
from playsound import playsound
import speech_recognition as sr

r = sr.Recognizer()

with sr.Microphone() as src:
    print(' Say Something')
    audio = r.listen(src)

try:
    t = r.recognize_google(audio, language='ar-AR' )
    print(t)
    f = open('text.txt','a',encoding='utf-8')
    f.writelines(t+'\n')
    f.close()
    obj = gTTS(text=t,lang='ar',slow=False)
    obj.save('text.mp3')
    playsound('text.mp3')

except sr.UnknownValueError as U:
    print(U)
except sr.RequestError as R:
    print(R)
```

Machine translation

Agenda

- Machine translation
- Important of machine translation
- Major issues involving Arabic
- Translation approaches
- Python code

Machine translation

- is a computer application that translates texts or speech from one natural language to another.
- Machine translation receives a source sentence,

  S = [s1, s2,...,si]

  and generates a target sentence,

  T = [t1 t2,..., tj]

  by translating the source sentence and give the meaning of it in the target language.

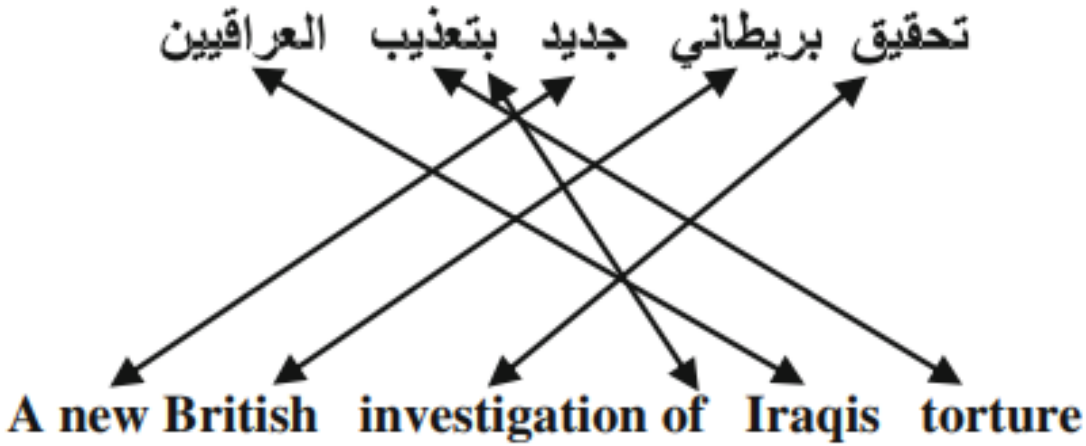
Important of machine translation

- Advent of computers, as there is an increasing demand to create online communication between people worldwide, speaking in different languages.
- Machine translation is a major administrative activity in natural language processing for different fields.

Online translation machines

- **Google Translator**, which is a free online text translation that is based on statistical machine translation paradigms and support more than 55 different languages.
- **Microsoft Translator** is based on example-based machine translation and several statistical machine translation technologies. It is a free online translation that supports 32 languages.
- **Systran** uses a rule based machine translation paradigm. Systran can translate a certain number of languages, like English, Arabic, French, Dutch, Chinese, and others.

Fig. 1 Arabic sentence and the equivalent sentence in English



Major issues involving Arabic

Arabic free word order

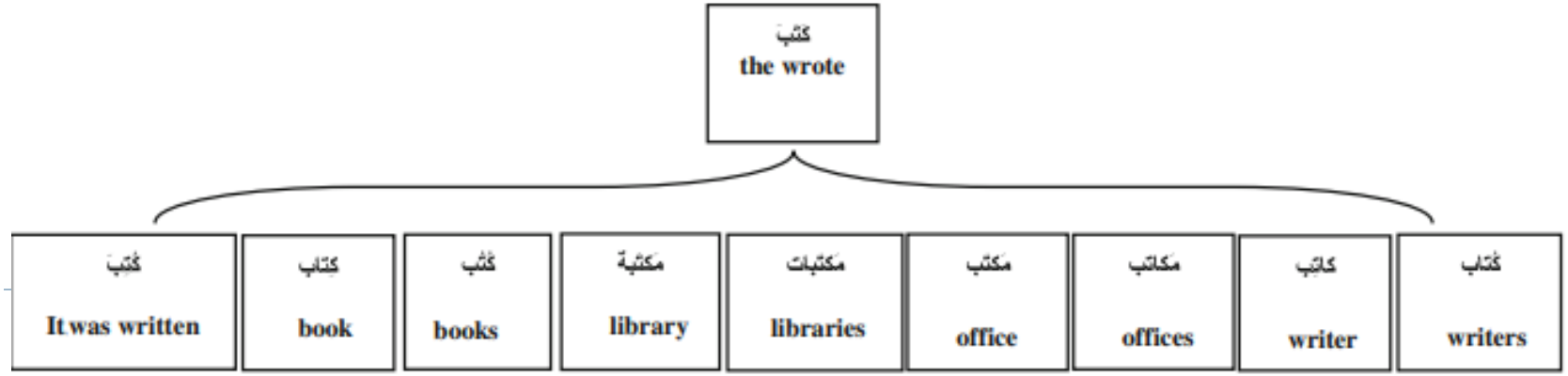
- Arabic has a different word order that provides a significant challenge to MT, due to the possibilities of expressing the same sentence in Arabic.
- In Arabic, three elements make-up a sentence, namely subject, verb, and object.
- Through these elements, Arabic can be classified into four types of sentences, according to different word orders i.e., SVO, VSO, VOS, and SOV.

Arabic free word order

Sentence orders	Arabic sentence	English translation
VSO	أكل ادم التفاحة	Ate Adam the apple
OVS	التفاحة أكل ادم	The apple ate Adam
SVO	ادم أكل التفاحة	Adam ate the apple
VOS	أكل التفاحة ادم	Ate the apple Adam

Derivation of words from a three-letter-root

- Arabic words can often be ambiguous, because of the three-letter root system.
- This system allows Arabic to develop to cover a wide choice of meanings.
- One or more of the root letters is dropped in some derivations, and this leads to possible ambiguity.



Arabic does not have copula verbs ‘to be’ and ‘to have’.

Arabic sentence	Arabic reading	English sentence(to be)
الباب مفتوح	The door opening	The door is opening
هو ذكي	He clever	He is clever

Arabic sentence	Arabic reading	English sentence(to have)
لها حقيبة	To her a bag	She has a bag
له كتاب	To him a book	He has a book

Feminine nouns are derived from masculine nouns

- Nouns in Arabic must either be masculine or feminine.
- Usually feminine nouns are derived from masculine nouns, which are considered as the stem.

Arabic nouns	English Translation
مهندس (male)	Engineer
مهندسة (female)	Engineer
طبيب (male)	Doctor
طبيبة (female)	Doctor

Feminine nouns are different from masculine

- Sometimes feminine nouns are different from masculine (the feminine nouns not derived from a masculine noun).

Arabic nouns	English translation
ولد (male)	Boy
بنت (female)	Girl
رجل (male)	Man
امرأة (female)	Woman
ديك (male)	Cock
دجاجة (female)	Chicken

Arabic dual and plural forms

- The number system in Arabic includes the dual form, whereas English moves from a singular to a plural form directly, but in Arabic we need to add a suffixing morpheme to the singular (stem) (or) depending on whether the case is nominative or accusative and genitive.

Arabic singular	Arabic dual (nominative)	Arabic dual (accusative and genitive)	English translation
مهندس (male) مهندسة (female)	مهندسان (male) مهندستان (female)	مهندسين (male) مهندستين (female)	Two engineers Two engineers

Plural form of Arabic Masculine nouns

- The plural form of Arabic masculine nouns exists by suffixing morpheme to the singular nouns (or) depending on whether the word case is nominative or accusative and genitive .

Arabic singular	Arabic plural (nominative)	Arabic plural (accusative and genitive)	English translation
معلم	معلمون	معلمين	Teachers
زائر	زائرون	زائرين	Visitors

Plural form of Arabic feminine nouns

- The plural form of Arabic feminine nouns can be created by adding a suffixing morpheme to the stem word (or) depending on whether the word case is nominative or accusative and genitive.

Arabic singular	Arabic plural (nominative)	Arabic plural (accusative and genitive)	English translation
معلمة	معلماتٌ	معلماتٍ	Teachers
زائرة	زائراتٌ	زائراتٍ	Visitors

Broken plural

- In Arabic, some words have no fixed rule for their plural form.
- Their plural forms are formed by changing the vowels, or adding or deleting the original alphabet; this type of plural is called a broken plural.

Arabic singular	English translation	Arabic plural	English translation
باب	Door	ابواب	Doors
قلم	Pen	اقلام	Pens
كوكب	Planet	كواكب	Planets

Translation approaches

- There are many different approaches to carrying out machine translation.
  - Rule-based
  - Statistical
  - Hybrid method

Rule-based

- Is the first technique used by researchers.
- Rules are **written by humans** according to their linguistic knowledge.
- The strength of this is that it can deeply analyze both syntax and semantic levels.
- In practice, rule-based machine translation systems often have diverse dictionaries, where some contain main entries, and others contain specialized vocabulary.

deep-translator

Python Example

deep-translator

- A flexible free and unlimited python tool to translate between different languages in a simple way using multiple translators.

```
pip install -U deep_translator
```

```
1 from deep_translator import GoogleTranslator
2
3 stm = GoogleTranslator(source='ar', target='en').translate('كيف حالك')
4 print(stm)
5
6 def T(text) :
7   return GoogleTranslator(source='ar', target='en').translate(text)
8
9 stm =T('ما عمرك')
10 print(stm)
```

References

- Khelifa, M.O., Elhadj, Y.M., Abdellah, Y. and Belkasmi, M., 2017. Constructing accurate and robust HMM/GMM models for an Arabic speech recognition system. *International Journal of Speech Technology*, *20*(4), pp.937-949.
- Al-Anzi, F. and AbuZeina, D., 2018, March. Literature survey of Arabic speech recognition. In *2018 International Conference on Computing Sciences and Engineering (ICCSE)* (pp. 1-6). IEEE.
- Al-Maadeed, N. and Al-Maadeed, S., 2018. Person-Dependent and Person-Independent Arabic Speech Recognition System. In *Recent Trends in Computer Applications* (pp. 267-278). Springer, Cham.

References

- Alqudsi, A., Omar, N. and Shaker, K., 2014. Arabic machine translation: a survey. *Artificial Intelligence Review*, *42*(4), pp.549-572.

Project Evaluation

Documentation	Definition	2
	Challenges of Arabic with this APP	3
	Related work (at least 3)	3
	future work	2
Implementation		5